

Построение и визуальный анализ параллельных решений для оптимизационных и параметрических исследований в вычислительной газовой динамике *

А.Е. Бондарев, В.А. Галактионов, Л.З. Шапиро

bond@keldysh.ru | vlgal@gin.keldysh.ru | pls1@gin.keldysh.ru

Москва, Россия, Институт прикладной математики им. М.В.Келдыша РАН

В работе рассматривается комбинированный подход, предназначенный для построения и анализа многомерных параметрических решений нестационарных задач вычислительной механики жидкости и газа. Подход может применяться для моделирования процессов образования пространственно-временных структур, а также для поиска оптимального сочетания геометрических или физических характеристик в выбранном классе задач.

Представленный в докладе подход основан на решении задач параметрического поиска и оптимизационного анализа с помощью параллельных вычислений. Результаты вычислений представляют собой многомерные массивы данных. Для поиска скрытых взаимозависимостей в массивах применяются методы анализа многомерных данных и визуализации. Все алгоритмы комбинированного подхода организованы в виде единой технологической цепочки. Подобная организация позволяет рассматривать реализованный подход как прототип обобщенного вычислительного эксперимента. Приводятся примеры практической реализации подхода.

Ключевые слова: многомерные параметрические решения, анализ многомерных данных, метод главных компонент

Design and Visual Analysis of Parallel Solutions for Optimizing and Parametric Studies in CFD *

A.E. Bondarev, V.A. Galaktionov, L.Z. Shapiro

Moscow, Russia, KIAM RAS

The paper presents a combined approach to design and analysis of multidimensional parametrical solutions for time-dependent CFD problems.

The approach can be applied to finding the conditions for space-time structures appearance and for the search of optimal combination of parameters (geometric or physical) for chosen class of problems. Presented approach is based on solutions of parametric studies and optimizing analysis problems by parallel computations.

The numerical results are presented in a form of multidimensional data volumes. To find out hidden dependencies in the volumes one should apply some methods for multidimensional data processing and visualizing. The algorithms of approach are organized in a form of pipeline. Such organization allows one to consider the approach as some kind of prototype for generalized numerical experiment. The examples of approach application to some practical problems are included.

Keywords: multidimensional parametric solutions, multidimensional data analysis, principal components

Введение

Современная вычислительная техника и численные методы позволяют в настоящее время тщательно и точно моделировать практически любой нестационарный физический процесс в механике сплошных сред и получать соответствующее поле физических величин. Однако в практических приложениях недостаточно просто рассчитать поле течения и представить красивую картину трансформации пространственно-временной структуры. Гораздо больший интерес вызывает не само явление, а то, при каких обстоятельствах оно возникает, т.е. зависимость возникновения явления от определя-

ющих параметров задачи, таких как числа Маха, Рейнольдса, Прандтля и т.д. Для того чтобы рассчитать подобную зависимость необходимо решение организовать решение задач параметрического поиска и оптимизационного анализа.

Параметрические численные исследования позволяют получать решение не для одной конкретной задачи математического моделирования, а для класса задач, заданного в многомерном пространстве определяющих параметров. Также применение параллельных алгоритмов на высокопроизводительной вычислительной технике позволяет численное исследование задач оптимизационного анализа, когда обратная задача решается в каждой точке сеточного разбиения многомерного пространства определяющих параметров. Основная особенность с точки зрения задач анализа и визуализации решений в подобных вычислениях за-

Работа выполнена при поддержке грантов Российского фонда фундаментальных исследований (проекты 16-01-00553а и 14-01-00769а). Работа опубликована при финансовой поддержке РФФИ, грант 16-07-20482.

ключается в том, что их результаты представляют собой многомерные массивы, размерность которых соответствует количеству определяющих параметров. Эти массивы нуждаются в обработке и визуальном представлении с целью их анализа и выявления внутренних взаимосвязей между определяющими параметрами. Подобные задачи начинают встречаться на практике все чаще, хотя следует отметить, что размерность подобных массивов на сегодняшний день ограничивается вычислительными мощностями и обычно составляет 4 – 5, в исключительных случаях – 6. Подробно вопросы постановки задач подобного типа и организации их численного решения рассмотрены в работе [1].

Задачи обработки, анализа и визуализации многомерных данных являются на сегодняшний день важным и актуальным направлением. Проблемы изучения многомерных объемов данных, задачи определения взаимного расположения точек в многомерном облаке данных, задачи выявления определяющих факторов и скрытых взаимосвязей между ними возникают практически во всех областях знания. Анализ многомерных данных (Data Analysis) интенсивно развивается как научная дисциплина, которая включает в себя: метод главных компонент (PCA-Principal Component Analysis) и его обобщения на нелинейные случаи, факторный, кластерный и дискриминантный анализ, построение самоорганизующихся карт (SOM – Self-Organized Maps) и упругих карт (Elastic Maps) [2]. В задачах вычислительной газовой динамики проблемы анализа многомерных данных ранее практически не встречались. Для обработки и визуального представления результатов даже самых сложных расчетов вполне хватало наработанных методов и приемов научной визуализации [4]. Однако в настоящее время интенсивное развитие высокопроизводительных и параллельных вычислений позволяет решать задачи параметрического исследования и задачи оптимизационного анализа [1].

В этой ситуации естественно хотелось бы применить уже наработанный аппарат методов и алгоритмов Data Analysis к подобным задачам, что вызывает необходимость адаптации методов Data Analysis для целей исследования многомерных результатов расчетов газодинамических задач.

Реализация параллельных алгоритмов

Общая постановка задачи оптимизационного анализа и вопросы оптимальной организации параллельных вычислений подробно описаны в [1]. Отмечалось, что в целом алгоритм решения задачи оптимизационного анализа сводится к решению большого количества обратных задач M^n , (при задании M точек в диапазоне разбиения каждого определяющего параметра), каждая из которых предполагает, в свою очередь решение большого количества

прямых задач. В данной ситуации необходимо применить параллельные вычисления. В качестве основного критерия пригодности для распараллеливания рассматривалась независимость от выбираемого для вычислений численного метода. Наиболее оптимальным и эффективным был признан вариант распараллеливания по принципу многозадачного параллелизма. При таком подходе вне зависимости от алгоритмов возможна организация параллельных вычислений однотипных обратных задач (ОЗ) с разными входными данными, представляющими собой фиксированные наборы определяющих параметров (ОП) по принципу «один вариант ОП – один процессор».

При реализации на многопроцессорной вычислительной системе общая схема параллельного варианта решения задачи оптимизационного анализа сводится к заданию разбиения по всем определяющим параметрам, формированию таким образом входных данных для однотипных обратных задач, заданию числа процессоров и раздаче заданий каждому процессору со своими входными параметрами. По завершении работы всех процессоров проводится сбор данных и формирование массива результатов для последующей обработки. В силу того, что процессы решения однотипных обратных задач происходят фактически без обменов информацией между процессорами, распараллеливание здесь сводится к организации интерфейса, управляющего распределением вариантов по процессорам и сбором данных в единый массив результатов. Данный вариант является наиболее легким в программной реализации и позволяет ускорить расчет во столько раз, сколько процессоров может быть выделено одновременно.

Для решения задач параметрического поиска и оптимизационного анализа были реализованы два параллельных интерфейса для MPI и DVM. Оба разработанных параллельных интерфейса работоспособны и могут применяться для широкого круга задач математического моделирования нестационарных процессов. Результатами вычислений подобного типа являются объемы многомерных данных, имеющие размерность, соответствующую числу определяющих параметров, и, как правило, больше трех.

Анализ и визуализация многомерных данных

В работе [4] рассматривались современные попытки построения визуальной концепции для представления многомерных данных, а также отмечалось отсутствие на сегодняшний день адекватного и надежного способа подобного визуального представления для объемов, имеющих размерность, превышающую 3. Следовательно, для анализа информации, содержащейся в полученном

многомерном массиве необходимо понизить размерность массива. Рассмотрим наиболее распространенные практические способы понижения размерности.

Рассматриваемые способы основаны на анализе дисперсий данных массива по координатным направлениям или нахождении в изучаемом многомерном пространстве вектора, по направлению которого дисперсия максимальна.

Первый способ представляет собой поиск координатного направления с наименьшей дисперсией. Вычисляются дисперсии по всем координатным направлениям, выбирается наименьшая из них, и в том случае, когда минимальная дисперсия существенно (на порядки) меньше остальных, значения исследуемой функции по координатному направлению с наименьшей дисперсией заменяются на константу, равную среднему значению по направлению. Таким образом, размерность исходного многомерного пространства понижается на единицу. Для пространств небольшой размерности $n = 4, 5$ во многих практических случаях данный подход работает успешно.

Второй распространенный способ понижения размерности заключается в построении графических проекций на стандартное число измерений $n \leq 3$ с фиксацией переменных, не участвующих в построении проекции.

Оба вышеизложенных подхода не являются строго обоснованными. Скорее, это алгоритмы выдвижения гипотез, нуждающихся в проверках. Однако эти методы позволяют получать реальные практические результаты.

Не менее эффективным является применение метода главных компонент (РСА). Суть метода состоит в переходе от исходной системы координат к новому ортогональному базису в рассматриваемом многомерном пространстве, оси которого ориентированы по направлениям максимальной дисперсии массива данных. Реализации метода главных компонент и алгоритмам его применения в различных областях посвящено большое количество литературы. Различные варианты реализации метода главных компонент и его обобщений для нелинейных случаев подробно представлены в [2]. Применение главных компонент дает нам возможность отобразить исследуемый многомерный массив на плоскость или в трехмерное пространство, образованное первыми тремя главными компонентами. В этом случае схема обработки, анализа и визуализации исходного многомерного объема данных будет выглядеть следующим образом. Для исходного объема вычисляются 3 первые главные компоненты, где каждая главная компонента является линейной комбинацией исходных переменных. Далее координаты исходных точек исследуемого объема

выражаются в координатах главных компонент и реализуется визуальное представление массива в двумерном или в трехмерном виде. Изучается полученное визуальное представление многомерного массива в главных компонентах и предпринимается попытка аппроксимации данных массива с помощью примитивных функций, имеющих аналитическое выражение.

Практические результаты

Данный приближенный подход был применен к многомерному объему данных, полученному, как результат решения задачи оптимизационного анализа нестационарного взаимодействия сверхзвукового потока вязкого сжимаемого теплопроводного газа со струйной преградой [1]. Преграда возникает благодаря недорасширенной спутной струе, истекающей из сопла, помещенного во внешний сверхзвуковой поток. При повышении скорости изменения степени нерасчетности струи возникает специфический режим течения, когда вещество струи распространяется вверх по потоку по внешней стенке сопла. Скорость изменения степени нерасчетности струи рассматривается как управляющий параметр задачи оптимизационного анализа. В качестве определяющих параметров задачи рассматривались характерные числа Маха, Рейнольдса, Прандтля и Струхала. Эти четыре параметра варьировались в определенных диапазонах. Целью решения задачи было нахождение скорости изменения степени нерасчетности струи, при которой реализуется специфический режим течения во всех диапазонах изменения характерных чисел задачи. В качестве результата решения задачи был получен 5-мерный объем данных, где в качестве переменных были 4 характерных числа задачи M_∞ , $\lg Re_\infty$, Pr , Sh_∞ и искомая скорость V^* . Для полученного многомерного объема были определены три первые главные компоненты. После перехода к главным компонентам строилось визуальное представление точек массива в главных компонентах (рис.1). Полученное визуальное представление многомерного массива в главных компонентах позволило предположить, что точки массива могут быть грубо аппроксимированы параметрически заданной плоскостью. После определения конкретного вида плоскости и ее коэффициентов было проведено обратное преобразование к исходным переменным и определение конкретного вида аппроксимирующей плоскости в исходных координатах. Это дало возможность получить искомую зависимость $V^* = F(M_\infty, \lg Re_\infty, Pr, Sh_\infty)$ в аналитическом виде. Полученные результаты представляют собой решение для класса задач, заданного в многомерном объеме определяющих параметров.) Другой пример практического применения комбинированного подхода приведен в работе [5], где рассмат-

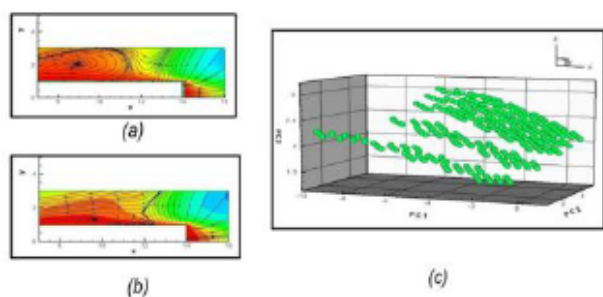


Рис. 1: Применение подхода к задаче взаимодействия потока со струйной преградой

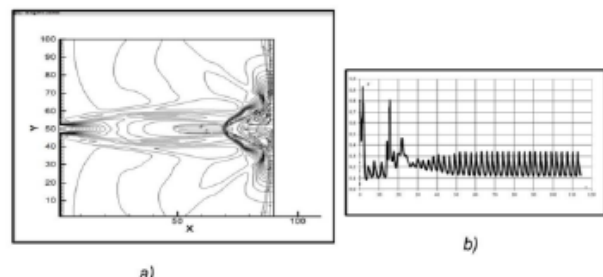


Рис. 2: Поиск пульсационных режимов при взаимодействии струи с поверхностью

ривается задача выявления условий возникновения осциллирующих режимов при взаимодействии сверхзвуковой недорасширенной струи с твердой поверхностью (рис.2). В качестве определяющих параметров задачи рассматривались три параметра – число Маха задачи M_∞ , степень нерасчетности струи n , соотношение удельных теплоемкостей γ . Эти параметры варьировались в определенных диапазонах. В качестве управляющего параметра выбиралось расстояние x от струи до поверхности. Целью решения задачи была проверка известного экспериментального соотношения между управляющим параметром и характерными параметрами задачи $x^2 = 16\gamma n M_\infty$. Результаты проведенных расчетов подтвердили факт подобной зависимости. Следующий пример применения подхода относит-

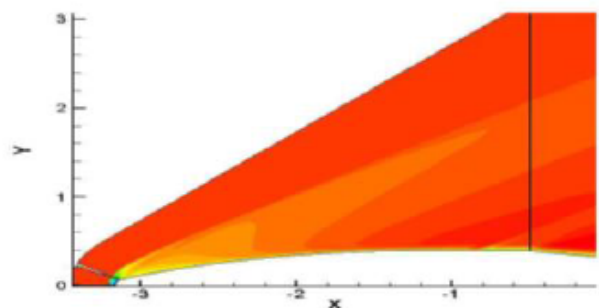


Рис. 3: Поле скорости при сверхзвуковом обтекании тела вращения

ся к параметрическим исследованиям. Рассматриваются нестационарные газодинамические процессы вокруг тела вращения в сверхзвуковом потоке (рис.3). Основной задачей является определение аэродинамических коэффициентов сопротивления тела вращения в зависимости от времени, скорости и вязкости набегающего потока, влажности, высоты полета, угла атаки. Чтобы рассчитать эти зависимости необходимо применять параллельные вычисления согласно вышеописанному подходу. В результате вычислений мы получаем многомерный объем данных. Далее он обрабатывается согласно вышеприведенному подходу, включая анализ дисперсий и применение метода главных компонент. В результате мы получаем искомые зависимости в квазианалитическом виде, что дает возможность использовать эти зависимости в дальнейших баллистических расчетах.

Обобщенный вычислительный эксперимент

Алгоритмы построения, обработки, анализа и визуального представления многомерных параметрических решений, будучи собранными воедино, представляют собой комплексный подход. Подход строится как единая технологическая цепочка алгоритмов производства, обработки, визуализации и анализа многомерных данных. Такая технологическая цепочка может рассматриваться как прототип обобщенного вычислительного эксперимента для нестационарных задач вычислительной газовой динамики. Схема реализации подобного обобщенного вычислительного эксперимента представлена на рисунке 4. Следует отметить, что подобный

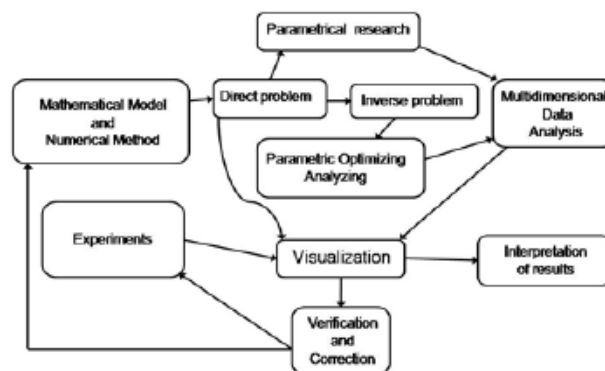


Рис. 4: Схема обобщенного вычислительного эксперимента

обобщенный вычислительный эксперимент неявно предполагает наличие надежной математической модели, численного метода для ее решения и набор экспериментальных результатов для верификации. В процессе вычислений необходимо реализовать организацию постоянного сравнения с экспе-

риментальными данными при наличии такой возможности. Набор используемых методов должен включать в себя решение обратных и оптимизационных задач. Будучи реализованными с помощью описанных ранее параллельных интерфейсов, эти методы позволят получать решения задач параметрического исследования и оптимизационного анализа в виде многомерных объемов данных. Для обработки этих объемов и выявления скрытых взаимозависимостей между изучаемыми в объеме параметрами необходимо интегрировать в общий набор алгоритмов методы анализа многомерных данных и их визуального представления. В итоге подобный обобщенный вычислительный эксперимент позволит получать решение не одной, отдельно взятой, задачи, а решение для целого класса задач, задаваемого диапазонами изменения определяющих параметров. Также следует отметить универсальность подобного обобщенного вычислительного эксперимента. Он может быть применен к широкому кругу задач математического моделирования нестационарных процессов. Практическая реализация подобного обобщенного эксперимента может обеспечивать организацию крупномасштабных промышленных расчетов.

Заключение

В работе рассмотрен комбинированный подход исследования условий возникновения пространственно-временных структур в нестационарных задачах вычислительной газовой динамики. Основная черта подхода – возможность сведения расчета к многократному решению одной задачи с различными наборами исходных параметров. Это позволяет применить параллельные вычисления и получить результаты в виде многомерных объемов данных. Для обработки и анализа этих объемов применяются методы визуальной аналитики. Все алгоритмы данного подхода выстраиваются в единую технологическую цепочку производства, обработки и анализа многомерных данных. Подобная технологическая цепочка может рассматриваться как прототип обобщенного вычислительного эксперимента.

Литература

- [1] Bondarev A.E., Galaktionov V.A. Parametric Optimizing Analysis of Unsteady Structures and Visualization of Multidimensional Data // International Journal of Modeling, Simulation and Scientific Computing, Vol. 4, suppl. issue 1, 2013, DOI: 10.1142/S1793962313410043 <http://www.worldscientific.com/doi/abs/10.1142/S1793962313410043>
- [2] Зиновьев А. Ю. Визуализация многомерных данных, Красноярск, Изд. КГТУ, 2000. 180 с.
- [3] J. Thomas and K. Cook, Illuminating the Path: Research and Development Agenda for Visual Analytics. IEEE-Press, 2005.
- [4] Бондарев А.Е., Галактионов В.А., Четкин В.М. Анализ развития концепций и методов визуального представления данных в задачах вычислительной физики / Журнал вычислительной математики и математической физики, 2011, Т. 51, N 4, С. 669–683.
- [5] Alexeev A.K., Bondarev A.E. Modeling of Time-Dependent Modes for Supersonic Underexpanded Jet Interacting with Plate // Mesh methods for boundary-value problems and applications. Proceedings of 10th International Conference. – Kazan: Otechestvo, 2014, pp.64-68.

Об авторах

Бондарев Александр Евгеньевич – к.ф.-м.н., старший научный сотрудник ИПМ им. М.В. Келдыша РАН E-mail: bond@keldysh.ru

Галактионов Владимир Александрович – д.ф.-м.н., профессор, заведующий отделом компьютерной графики и вычислительной оптики ИПМ им. М.В. Келдыша РАН E-mail: vlgal@gin.keldysh.ru

Шапиро Лев Залманович – к.ф.-м.н., старший научный сотрудник ИПМ им. М.В. Келдыша РАН E-mail: pls@gin.keldysh.ru